

5th
AN

1. A method of correcting incorrect text associated with recognition errors in computer-implemented speech recognition, the method comprising:

- performing speech recognition on an utterance to produce a recognition result for the utterance;
- receiving a selection of a word from the recognized utterance, the selection indicating a bound of a portion of the recognized utterance to be corrected;
- comparing a first alternative transcript to the recognized utterance to be corrected;
- producing a first recognition correction based on the comparison;
- comparing a second alternative transcript to the recognized utterance to be corrected;
- producing a second recognition correction based on the second comparison; and
- replacing a portion of the recognition result with one of the first recognition correction and the second recognition correction;

wherein a duration of the first recognition correction differs from a duration of the second recognition correction, and the portion of the recognition result replaced includes at one bound a word indicated by the selection and extends for the duration of the one of the first recognition correction and the second recognition correction with which the portion is replaced.

2. The method of claim 1 wherein the selection indicates a beginning bound of a recognized utterance to be corrected.

3. The method of claim 1 wherein the selection indicates a finishing bound of a recognized utterance to be corrected.

4. The method of claim 1 wherein comparing an alternative transcript to the recognized utterance comprises:

- selecting from the alternative transcript a test word that is not identical to the selected word and that begins at a time that is nearest a time at which the selected word begins; and
- searching in time, through the recognized utterance relative to the selected word, and through the alternative transcript relative to the test word, until finding a word common to the recognized utterance and the alternative transcript.

1 5. The method of claim 4 wherein the common word begins at a time in the
2 recognized utterance that is approximately equal to a time at which the common word begins
3 in the alternative transcript.

1 6. The method of claim 5 wherein producing a recognition correction comprises
2 selecting a word string from the alternative transcript, the word string bound by the test word
3 from the alternative transcript and by a word from the alternative transcript that is adjacent to
4 the common word and between the test word and the common word.

1 7. The method of claim 6 further comprising receiving a selection of one of the
2 first recognition correction and the second recognition correction.

1 8. The method of claim 6 wherein searching in time through the recognized
2 utterance and through the alternative transcript comprises:
3 designating a word adjacent to the test word as an alternative transcript word;
4 designating a word adjacent to the selected word as an original transcript word; and
5 comparing the original transcript word to the alternative transcript word.

1 9. The method of claim 8 wherein searching in time through the recognized
2 utterance and through the alternative transcript comprises:

3 designating the original transcript word and the alternative transcript word as the
4 common word if:

5 the original transcript word is identical to the alternative transcript word, and
6 a time at which the original transcript word begins is near a time at which the
7 alternative transcript word begins.

1 10. The method of claim 9 wherein searching in time through the recognized
2 utterance and through the alternative transcript comprises:

3 designating a word adjacent to the alternative transcript word in the alternative
4 transcript as the alternative transcript word if:

5 the original transcript word is not identical to the alternative transcript word,
6 and

7 a time at which the original transcript word begins is later than a time at which
8 the alternative transcript word begins.

1 11. The method of claim 9 wherein searching in time through the recognized
2 utterance and through the alternative transcript comprises:

3 designating a word adjacent to the alternative transcript word in the alternative
4 transcript as the alternative transcript word if:

5 the original transcript word is identical to the alternative transcript word, and
6 a time at which the original transcript word begins is later than a time at which
7 the alternative transcript word begins.

1 12. The method of claim 9 wherein searching in time through the recognized
2 utterance and through the alternative transcript comprises:

3 designating a word adjacent to the original transcript word in the original transcript as
4 the original transcript word if:

5 the original transcript word is not identical to the alternative transcript word,
6 and

7 a time at which the original transcript word begins is earlier than a time at
8 which the alternative transcript word begins.

1 13. The method of claim 9 wherein searching in time through the recognized
2 utterance and through the alternative transcript comprises:

3 designating a word adjacent to the original transcript word in the original transcript as
4 the original transcript word if:

5 the original transcript word is identical to the alternative transcript word, and
6 a time at which the original transcript word begins is earlier than a time at
7 which the alternative transcript word begins.

1 14. The method of claim 9 wherein searching in time through the recognized
2 utterance and through the alternative transcript comprises:

3 designating a word adjacent to the original transcript word in the original transcript as
4 the original transcript word and designating a word adjacent to the alternative transcript word
5 in the alternative transcript as the alternative transcript word if:

the original transcript word is not identical to the alternative transcript word,
and
a time at which the original transcript word begins is near a time at which the
alternative transcript word begins.

15. A method of correcting incorrect text associated with recognition errors in
computer-implemented speech recognition, the method comprising:
performing speech recognition on an utterance to produce a recognition result for the
utterance;
receiving a selection of a word from the recognized utterance, the selection indicating
a bound of a portion of the recognized utterance to be corrected;
comparing an alternative transcript to the recognized utterance to be corrected, the
comparing comprising:
selecting from the alternative transcript a test word that begins at a time that is
nearest a time at which the selected word begins; and
searching in time, relative to the selected word, through the recognized
utterance, and searching in time, relative to the test word, through the alternative transcript
until finding a word common to the recognized utterance and the alternative transcript;
producing a recognition correction based on the comparison; and
replacing a portion of the recognition result with the recognition correction.

16. A method of correcting incorrect text associated with recognition errors in
computer-implemented speech recognition, the method comprising:
performing speech recognition on an utterance to produce a recognition result for the
utterance;
receiving a selection of a word from the recognized utterance, the selection indicating
a bound of a portion of the recognized utterance to be corrected; and
comparing an alternative transcript to the recognized utterance to be corrected, the
comparing including:
selecting from the alternative transcript a test word that is not identical to the
selected word and that occurs at a time that is nearest a time at which the selected word
occurs, and

12 searching in time, through the recognized utterance relative to the selected
13 word, and through the alternative transcript relative to the test word, until finding a word
14 common to the recognized utterance and the alternative transcript.

1 17. The method of claim 16, further comprising:
2 producing a recognition correction based on the comparison;
3 replacing a portion of the recognition result with the recognition correction;
4 wherein the portion of the recognition result replaced includes at one bound a word
5 indicated by the selection and extends for the duration of the recognition correction with
6 which the portion is replaced.

1 18. A method of correcting incorrect text associated with recognition errors in
2 computer-implemented speech recognition, the method comprising:
3 receiving a text document formed by recognizing speech utterances using a
4 vocabulary;
5 receiving a general confusability matrix having one or more values each indicating a
6 likelihood of confusion between a first phoneme and a second phoneme;
7 receiving corrected text that corresponds to misrecognized text from the text
8 document;
9 generating a sequence of phonemes for the corrected text;
10 aligning the generated sequence of phonemes with phonemes of the misrecognized
11 text;
12 adjusting one or more values of the general confusability matrix based on the
13 alignment to form a specific confusability matrix; and
14 searching the text document for additional instances of the corrected text using the
15 specific confusability matrix.

1 19. The method of claim 18 in which generating the sequence of phonemes for the
2 corrected text includes determining whether the corrected text is in the vocabulary.

1 20. The method of claim 18 further comprising outputting the text document.

1 21. The method of claim 18 in which a list of recognition candidates is associated
2 with each recognized speech utterance.

1 22. The method of claim 18 in which generating the sequence of phonemes for the
2 corrected text comprises using a phonetic alphabet.

1 23. The method of claim 18 further comprising generating the general
2 confusability matrix using empirical data.

1 24. The method of claim 23 in which the empirical data comprises information
2 relating to a rate of confusion of phonemes for a preselected population.

1 25. The method of claim 21 in which the empirical data comprises information
2 relating to frequency characteristics of different phonemes.

1 26. The method of claim 21 in which the empirical data comprises information
2 acquired during an adaptive training of a user.

1 27. The method of claim 18 in which searching the text document for the
2 corrected text comprises searching the text document for the sequence of phonemes for the
3 corrected text.

1 28. The method of claim 18 in which searching the text document for the
2 corrected text comprises searching the text document for a sequence of phonemes that is
3 likely to be confused with the sequence of phonemes for the corrected text.

1 29. The method of claim 18 in which searching the text document for the
2 corrected text comprises scoring a portion of the text document and comparing the score of
3 the portion to an empirically determined threshold value to determine whether the portion of
4 the text document includes a word that is not in the vocabulary.

1 30. The method of claim 29 further comprising outputting a result if it is
2 determined that the portion of the text document includes a word that is not in the
3 vocabulary.

1 31. The method of claim 30 in which outputting the result comprises highlighting
2 the portion of the text document.

1 32. The method of claim 30 in which outputting the result comprises re-
2 recognizing the portion of the text document.

1 33. A computer-implemented method for speech recognition, the method
2 comprising:

3 receiving dictated text;

4 generating recognized speech based on the received dictated text, the generating
5 comprising determining acoustic models for the dictated text that best match acoustic data for
6 the dictated text;

7 receiving an edited text of the recognized speech, the edited text indicating a
8 replacement for a portion of the dictated text;

9 determining an acoustic model for the edited text;

10 determining whether to adapt acoustic models for the edited text based on the
11 acoustic model for the edited text and the acoustic model for the dictated text portion.

1 34. The method of claim 33 further comprising calculating an acoustic model
2 score based on a comparison between the acoustic model for the edited text and the acoustic
3 data for the dictated text portion.

1 35. The method of claim 34 in which determining whether to adapt acoustic
2 models for the edited text is based on the calculated acoustic model score.

1 36. The method of claim 35 in which determining whether to adapt acoustic
2 models for the edited text comprises calculating an original acoustic model score based on a
3 comparison between the acoustic model for the dictated text portion and the acoustic data for
4 the dictated text portion.

1 37. The method of claim 36 in which determining whether to adapt acoustic
2 models for the edited text comprises calculating a difference between the acoustic model
3 score and the original acoustic model score.

1 38. The method of claim 37 in which determining whether to adapt acoustic
2 models for the edited text comprises determining whether the difference is less than a
3 predetermined value.

1 39. The method of claim 38 in which determining whether to adapt acoustic
2 models for the edited text comprises adapting acoustic models for the edited text if the
3 difference is less than a predetermined value.

1 40. The method of claim 38 in which determining whether to adapt acoustic
2 models for the edited text comprises bypassing adapting acoustic models for the edited text if
3 the difference is greater than or equal to a predetermined value.

1 41. The method of claim 33 in which receiving the edited text of the recognized
2 speech occurs during a recognition session in which the recognized speech is generated.

1 42. The method of claim 33 in which receiving the edited text of the recognized
2 speech occurs after a recognition session in which the recognized speech is generated.

1 43. The method of claim 33 in which receiving the edited text of the recognized
2 speech comprises receiving a selection of the portion of the dictated text.

1 44. The method of claim 33 in which determining an acoustic model for the edited
2 text comprises searching for the edited text in a vocabulary or a backup dictionary used to
3 generate the recognized speech.

1 45. The method of claim 33 in which determining an acoustic model for the edited
2 text comprises selecting an acoustic model that best matches the edited text.

1 46. A computer-implemented method of speech recognition, the method
2 comprising:
3 performing speech recognition on an utterance to produce a recognition result for the
4 utterance;
5 receiving a selection of the recognition result;
6 receiving a correction of the recognition result;

7 performing speech recognition on the correction using a constraint grammar that
8 permits spelling and pronunciation in parallel; and
9 identifying whether the correction comprises a spelling or a pronunciation using the
10 constraint grammar.

1 47. The method of claim 46 further comprising generating a replacement result
2 for the recognition result based on the correction.

1 48. The method of claim 46 in which the constraint grammar includes a spelling
2 portion and a dictation vocabulary portion.

1 49. The method of claim 48 in which the spelling portion indicates that the first
2 utterance from the user is a letter in an alphabet.

1 50. The method of claim 48 in which the vocabulary portion indicates that the
2 first utterance from the user is a word from the dictation vocabulary.

1 51. The method of claim 48 in which the spelling portion indicates a frequency
2 with which letters occur in a language model.

1 52. The method of claim 48 in which the dictation vocabulary portion indicates a
2 frequency with which words occur in a language model.

1 53. The method of claim 48 further comprising introducing a biasing value
2 between the spelling and the dictation vocabulary portions of the constraint grammar.